Supplementary Materials

Supplementary Table 1. Exclusions performed in the genome-wide association study (GWAS) patients and independent replication control subjects

Reason for exclusion	GWAS case patients, No. (%)	Independent replication series, No. (%)	
		Case patients	Control subjects
Total samples genotyped	1238 (100)	599 (100)	633 (100)
Completely failed	6 (0.48)	10 (1.67)	6 (0.95)
Genotype rate	3 (0.24)	23 (3.84)	13 (2.05)
Autosomal heterozygosity rate	6 (0.48)	NA	NA
Sex discrepancy	5 (0.40)	3 (0.50)	1 (0.16)
Relatedness (first-degree relatives)	5 (0.40)	NA	NA
Ethnic heterogeneity based on PCA*	13 (1.05)	NA	NA
Total included in analysis	1200 (96.9)	563 (93.9)	613 (96.8)

* Principal components analysis (PCA) was conducted using a subseries of 11029 single-nucleotide polymorphisms present on all Illumina BeadChip panels (Illumina Inc., San Diego, CA) that are evenly distributed across the genome in low linkage disequilibrium. NA = not applicable.

Supplementary Table 2. Comparison between genome-wide association study (GWAS) genotypes and technical replication genotyping in case patients and study-specific control subjects for select single-nucleotide polymorphism (SNPs)

Chromosome	SNP	Allele		Case patients		Control subjects	
				Frequency*	Genotype Concordance†	Frequency*	
		A1*	A2	Sequenom	No. of case patients (%)	Generic	Study-specific
5	rs20541	A	G	0.24	1210 (100)	0.19	0.20
6	rs9393777‡	C	T	NA	NA	0.14	NA
6	rs3799499	T	G	0.28	1210 (100)	0.25	0.23
6	rs2523399	C	T	0.50	1151 (100)	0.46	0.43
6	rs2734986‡	C	T	NA	NA	0.17	NA
6	rs6904029	A	G	0.27	1207 (99.5)	0.30	0.31
6	rs1150741	G	C	0.34	NA	NA	0.30
6	rs1245371‡	G	A	NA	NA	0.28	NA
6	rs1049623	C	T	0.40	1205 (99.9)	0.39	0.38
6	rs3094211	G	A	0.30	1209 (100)	0.26	0.26
6	rs3094204	G	A	0.45	1197 (100)	0.49	0.52
6	rs746647	G	A	0.35	1208 (100)	0.31	0.31
6	rs3130542	A	G	0.24	1205 (100)	0.20	0.21
6	rs2248462	A	G	0.14	1204 (100)	0.22	0.20
6	rs206015	A	G	0.07	1208 (100)	0.11	0.12
6	rs443198	G	A	0.32	1172 (99.0)	0.36	0.37
6	rs3134931	C	T	0.22	1177 (99.8)	0.30	0.32
6	rs411326	T	C	0.19	1201 (100)	0.26	0.25
6	rs2395174	G	T	0.31	1207 (100)	0.30	0.29
6	rs6903608	C	T	0.41	1210 (100)	0.31	0.30
6	rs2395185	T	G	0.22	1204 (100)	0.32	0.32
6	rs7775228	C	T	0.08	1210 (100)	0.12	0.12
6	rs9275572	A	G	0.54	1187 (100)	0.44	0.43

* Frequency of the minor allele (A1). NA = not applicable.

† Percent concordance among subjects for which genotypes were available from both the GWAS and technical replication Sequenom genotyping (Sequenom Inc., San Diego, CA).

‡ These SNPs were not successfully genotyped in the technical replication stage.

Supplementary Table 3. Characteristics of classical Hodgkin lymphoma case patients and control subjects

Characteristics	Genome-wide association study			Independent replication	
	Case patients, No. (%)	Generic control subjects, No. (%)	Study-specific control subjects, No. (%)	Case patients, No. (%)	Control subjects, No. (%)
Total	1200 (100)	6417 (100)	1395 (100)	563 (100)	613 (100)
Country					
Czech Republic	54 (4.50)	543 (8.46)	106 (7.60)	2 (0.36)	21 (3.43)
Denmark	189 (15.8)	81 (1.26)	525 (37.6)	0 (0)	0 (0)
France	13 (1.08)	157 (2.45)	21 (1.51)	16 (2.84)	49 (7.99)
Germany	86 (7.17)	52 (0.81)	187 (13.4)	6 (1.07)	86 (14.03)
Ireland	3 (0.25)	12 (0.19)	5 (0.36)	15 (2.66)	26 (4.24)
Italy	0 (0)	0 (0)	0 (0)	3 (0.53)	20 (3.26)
Netherlands	287 (23.9)	1804 (28.1)	0 (0)	0 (0)	0 (0)
Norway	0 (0)	389 (6.06)	0 (0)	0 (0)	0 (0)
Spain	25 (2.08)	99 (1.54)	58 (4.16)	22 (3.91)	194 (31.7)
Sweden	152 (12.7)	108 (1.68)	145 (10.4)	0 (0)	0 (0)
United Kingdom	391 (32.6)	3172 (49.4)	348 (25.0)	499 (88.6)	217 (35.4)
Gender					
Male	637 (53.1)	3489 (54.4)	795 (57.0)	321 (57.0)	334 (54.5)
Female	563 (46.9)	2928 (45.6)	600 (43.0)	242 (43.0)	279 (45.5)
Age, y					
15−24	295 (24.6)	6 (0.09)	135 (9.68)	153 (27.2)	57 (9.30)
25−34	332 (27.7)	28 (0.44)	181 (13.0)	147 (26.1)	119 (19.4)
35−44	210 (17.5)	217 (3.38)	189 (13.6)	94 (16.7)	123 (20.1)
45−54	137 (11.4)	3526 (55.0)	200 (14.3)	63 (11.2)	122 (19.9)
55−64	126 (10.5)	1159 (18.1)	306 (21.9)	68 (12.1)	101 (16.5)
65−74	86 (7.17)	1175 (18.3)	362 (26.0)	27 (4.80)	76 (12.4)
75−84	12 (1.00)	289 (4.50)	19 (1.36)	9 (1.60)	14 (2.28)
85−94	0 (0)	11 (0.17)	2 (0.14)	1 (0.18)	1 (0.16)
missing	2	6	1	1	0

Supplementary Table 4. Summary of classical Hodgkin lymphoma (cHL) results for single-nucleotide polymorphisms (SNPs) recently reported in genome-wide association studies (GWAS) of psoriasis and asthma *

Chromosome	SNP	Position	Gene	Total cHL		EBV+ cHL		EBV- cHL		Previous GWAS phenotype†
				OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	
5q31	rs20541	132023863	*IL13*	1.43 (1.28 to 1.60)	6.0×10^{-10}	1.07 (0.85 to 1.35)	.56	1.53 (1.32 to 1.76)	5.4×10^{-9}	Asthma, psoriasis, serum IgE
5q15	rs27524	96127700	*ERAP1*	1.21 (1.10 to 1.33)	1.5×10^{-4}	1.27 (1.05 to 1.53)	.01	1.24 (1.10 to 1.40)	5.5×10^{-4}	Psoriasis
1p13	rs2476601	114179091	*PTPN22*	1.24 (1.07 to 1.44)	3.7×10^{-3}	1.23 (0.92 to 1.64)	.16	1.26 (1.05 to 1.51)	.02	Psoriasis
2p16	rs702873	60935046	*REL*	1.12 (1.02 to 1.24)	.02	1.06 (0.88 to 1.27)	.54	1.15 (1.02 to 1.30)	.02	Psoriasis
19p13	rs12720356	10330975	*TYK2*	1.18 (1.01 to 1.39)	.04	1.30 (0.96 to 1.78)	.09	1.19 (0.97 to 1.45)	.10	Psoriasis
1p31	rs2201841	67466790	*IL23R*	1.10 (0.99 to 1.21)	.08	0.95 (0.78 to 1.16)	.62	1.15 (1.01 to 1.31)	.03	Psoriasis
1p36	rs4649203	24392507	*IL28RA*	0.94 (0.85 to 1.05)	.27	1.04 (0.85 to 1.28)	.71	0.88 (0.76 to 1.01)	.06	Psoriasis
5q33	rs2546890	158692478	*IL12B*	0.95 (0.87 to 1.05)	.31	0.92 (0.77 to 1.11)	.40	0.94 (0.84 to 1.07)	.35	Psoriasis
7q36	rs916514	154099909	*DPP6*	1.07 (0.92 to 1.25)	.40	1.01 (0.75 to 1.37)	.93	1.09 (0.89 to 1.32)	.40	Psoriasis
12q13	rs2066808	55024240	*STAT2*	1.06 (0.88 to 1.27)	.54	1.04 (0.73 to 1.49)	.82	1.04 (0.82 to 1.32)	.74	Psoriasis
11q22	rs1939015	102081585	*MMP27*	1.04 (0.91 to 1.18)	.58	1.00 (0.77 to 1.30)	.99	1.08 (0.92 to 1.27)	.36	Psoriasis
5q33	rs953861	158705160	*IL12B*	1.03 (0.91 to 1.17)	.63	1.05 (0.83 to 1.33)	.70	1.05 (0.90 to 1.22)	.55	Psoriasis
16p12	rs1859308	27305499	*IL21R*	1.03 (0.90 to 1.18)	.67	1.23 (0.95 to 1.58)	.11	0.97 (0.81 to 1.16)	.74	Serum IgE
15q22	rs744910	65233839	*SMAD3*	0.98 (0.89 to 1.08)	.70	1.07 (0.89 to 1.29)	.44	0.93 (0.83 to 1.05)	.27	Asthma
6p21	rs240993	111780407	*REV3L*	1.02 (0.92 to 1.14)	.70	1.00 (0.81 to 1.24)	1.0	1.04 (0.90 to 1.19)	.62	Psoriasis
9p24	rs1342326	6180076	*IL33*	0.98 (0.86 to 1.11)	.73	1.07 (0.84 to 1.36)	.61	0.93 (0.79 to 1.10)	.42	Asthma
17q12	rs2305480	35315722	*GSDML*	1.02 (0.93 to 1.12)	.73	0.92 (0.76 to 1.10)	.35	1.02 (0.91 to 1.16)	.70	Asthma
12q13	rs167769	55790042	*STAT6*	0.98 (0.89 to 1.08)	.74	1.05 (0.88 to 1.27)	.58	0.96 (0.84 to 1.08)	.47	Serum IgE
6p21	rs458017	111802784	*REV3L*	0.97 (0.80 to 1.19)	.80	1.08 (0.74 to 1.57)	.71	0.97 (0.75 to 1.25)	.80	Psoriasis
6p21	rs10484554	31382534	*HLA-C*	1.01 (0.87 to 1.16)	.92	1.35 (1.05 to 1.74)	.02	0.90 (0.75 to 1.10)	.30	Psoriasis
22q12	rs2284033	35863980	*IL2RB*	1.00 (0.91 to 1.10)	.99	1.03 (0.85 to 1.24)	.77	0.98 (0.87 to 1.11)	.78	Asthma

*Odds ratio (OR) and 95% confidence interval (CI) were derived using logistic regression assuming a log-additive genetic model of inheritance and adjusting for sex (male or female), country (eight indicator variables after excluding one country as the reference), and eight principal components analysis eigenvectors. All statistical tests were two-sided. EBV = Epstein-Barr virus. †Indicates the phenotype that the SNP was associated with in recent genome-wide association studies of psoriasis (30, 31) and asthma/immunoglobulin E (IgE) (28, 29).

Supplementary Table 5. Summary of genome-wide association study (GWAS) and technical replication stage results for the five independently associated major histocompatibility complex (MHC) region single-nucleotide polymorphisms (SNPs) and rs20541 (5q31, *IL13*), excluding the Netherlands subjects *

SNP (region)	Gene candidate	GWAS†		Technical replication†		P_{hom}‡
		OR (95% CI)	*P*	OR (95% CI)	*P*	
Total cHL						
rs2248462 (6p21)	*MICB*	0.64 (0.55 to 0.75)	1.8×10^{-8}	0.65 (0.55 to 0.77)	5.0×10^{-7}	.89
rs2395185 (6p21)	*HLA-DRA*	0.54 (0.47 to 0.61)	2.2×10^{-20}	0.57 (0.49 to 0.66)	1.3×10^{-14}	.59
EBV+ cHL						
rs2734986 (6p21) §	*HLA-A*	2.31 (1.78 to 2.99)	2.5×10^{-10}	NA	NA	NA
rs6904029 (6p21)	*HCG9*	0.51 (0.39 to 0.68)	2.6×10^{-6}	0.50 (0.38 to 0.66)	1.1×10^{-6}	.92
EBV- cHL						
rs6903608 (6p21)	*HLA-DRA*	2.13 (1.82 to 2.48)	2.7×10^{-21}	1.98 (1.67 to 2.35)	5.4×10^{-15}	.54
rs20541 (5q31)	*IL13*	1.48 (1.23 to 1.77)	2.4×10^{-5}	1.40 (1.15 to 1.70)	8.7×10^{-4}	.68

* Study-specific control subjects for the Northern Dutch Hodgkin Lymphoma Study were not available for the technical replication stage. To make results directly comparable, the technical replication analysis excluded the Netherlands case patients, and the GWAS analysis excluded the Netherlands case patients and control subjects. cHL = classical Hodgkin lymphoma, CI = confidence interval, EBV = Epstein-Barr virus, NA = not applicable, OR = odds ratio.

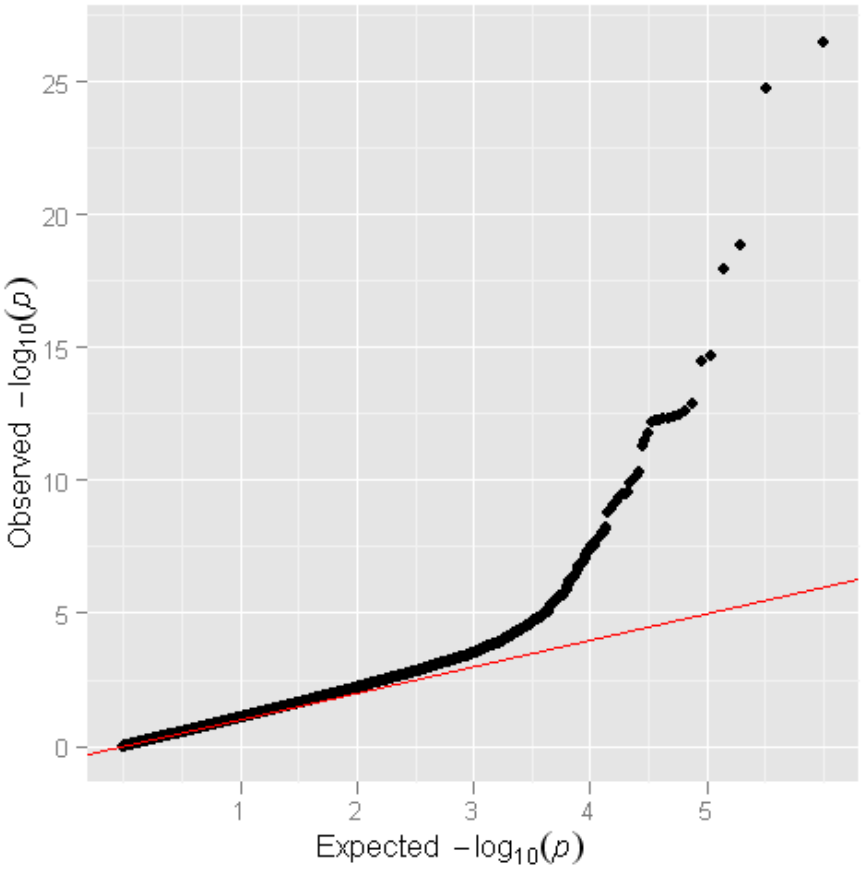
† The GWAS and technical replication analyses included 883 case patients (194 EBV-positive and 448 EBV-negative cHL). The GWAS and technical replication analysis included 4613 generic and 1395 study-specific control subjects, respectively. ORs and 95% CIs were derived using logistic regression assuming a log-additive genetic model of inheritance and adjusting for sex (male or female)

and country (seven indicator variables after excluding one country as the reference). The genome-wide association analysis additionally included eight principal components analysis eigenvectors. All statistical tests were two-sided.

‡ Two-sided *P* value for the χ^2 test of homogeneity (P_{hom}) comparing the risk estimates derived from the GWAS and technical replication analyses.

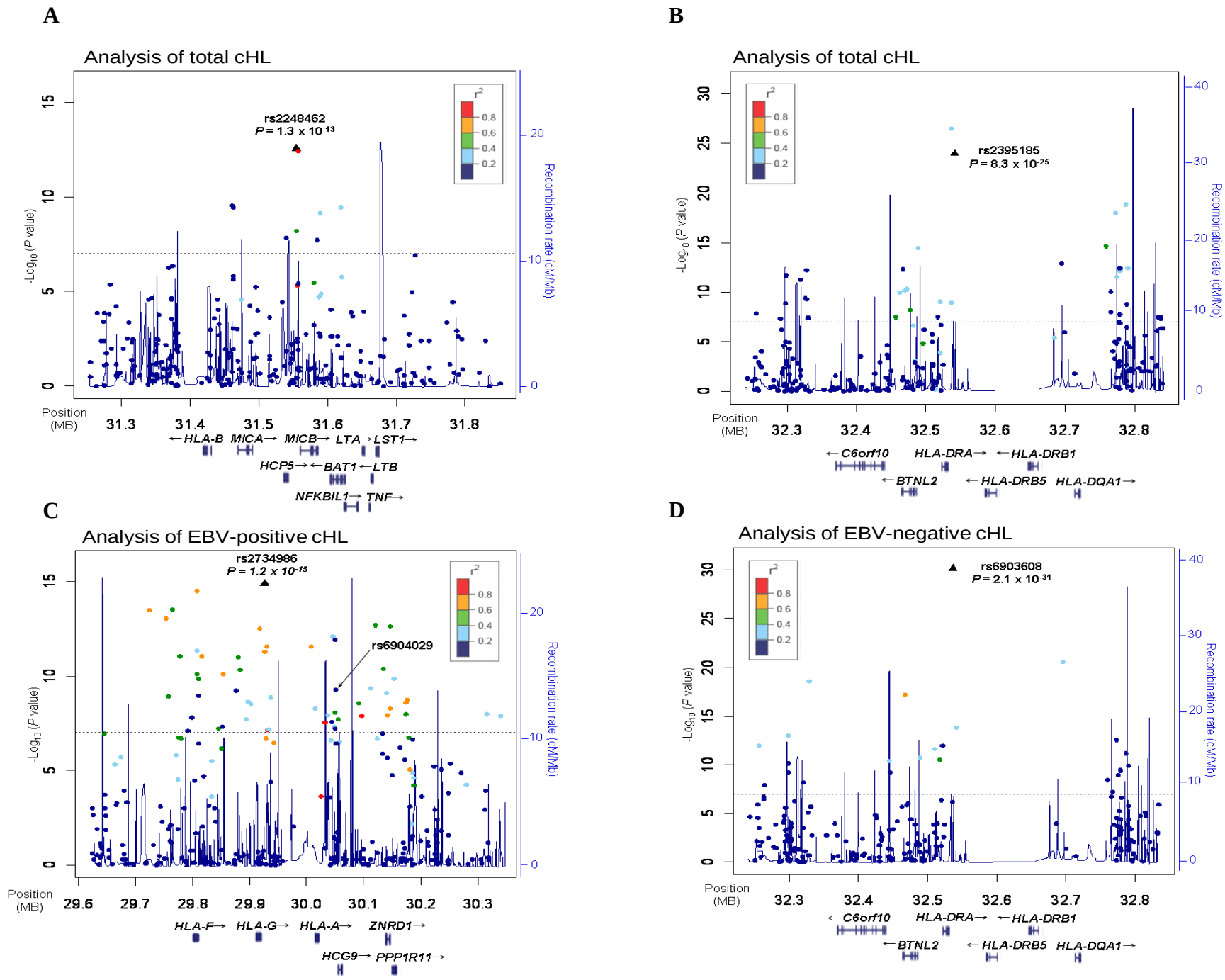
§ This SNP was not successfully genotyped in the technical replication stage.

Supplementary Figures

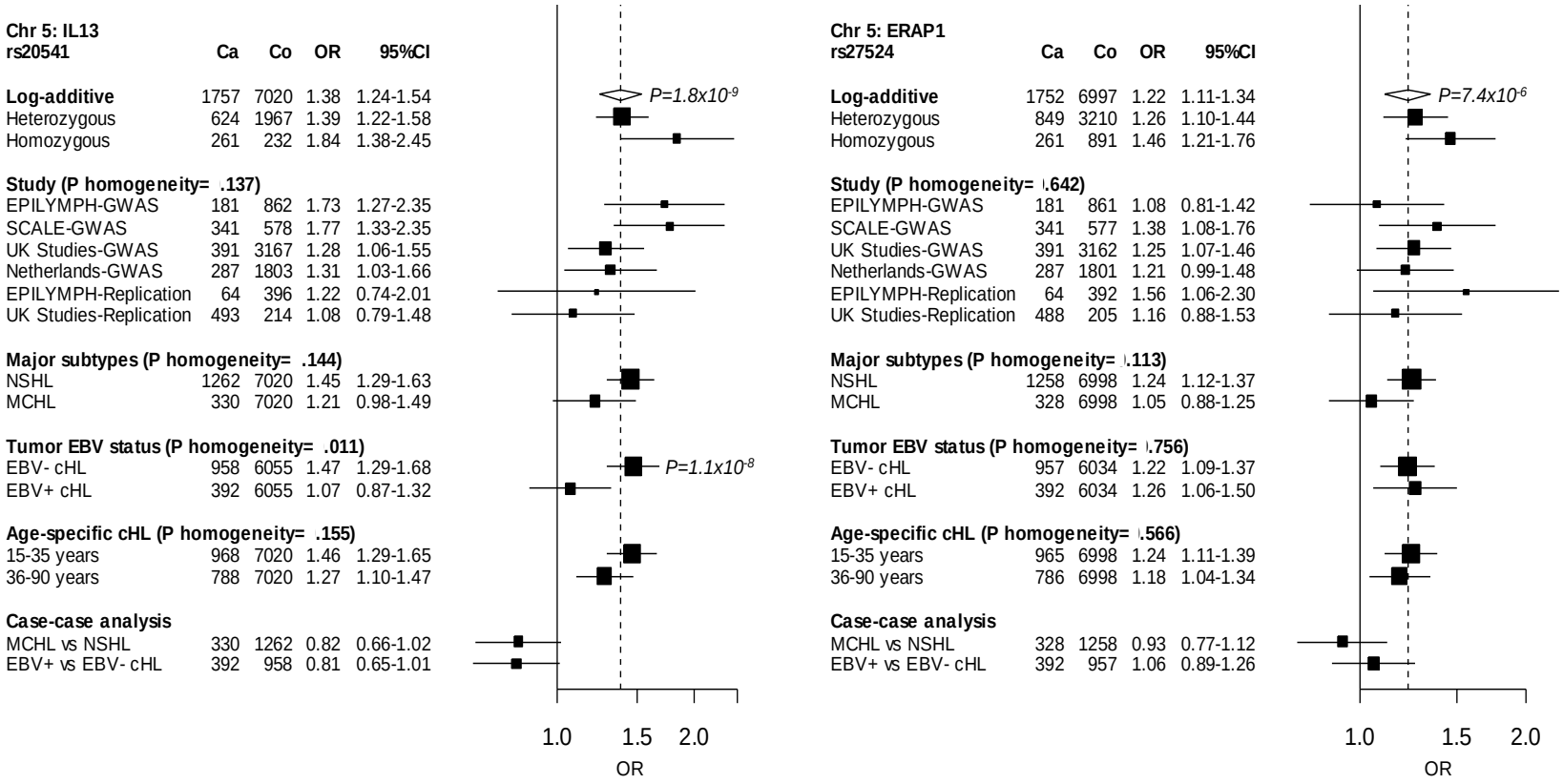


Supplementary Figure 1. Quantile-quantile plot of the expected vs observed -$\log_{10}$ (*P* value) distribution in the genome-wide association analysis of total classical Hodgkin lymphoma. Association results were derived by multiple logistic regression assuming a log-additive genetic model and adjusting for sex (male or female), country (France, Germany, Spain, Czech Republic, Ireland, United

Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. The **red line** represents the plot where the observed distribution of the $-\log_{10}$ (*P* value) is same as the expected distribution given the number of SNPs tested. The genomic inflation factor adjusted to the sample size of 1000 case patients and 1000 control subjects, using the method of de Bakker et al. (39), was equal to 1.04. All statistical tests were two-sided.


A
Analysis of total cHL
rs2248462
P = 1.3 x 10-13
r2
0.8
0.6
0.4
0.2
-Log10 (P value)
Recombination rate (cM/Mb)
Position (MB)
31.3
31.4
31.5
31.6
31.7
31.8
HLA-B
MICA
MICB
LTA
LST1
HCP5
BAT1
LTB
NFKBIL1
TNF
B
Analysis of total cHL
rs2395185
P = 8.3 x 10-25
32.3
32.4
32.5
32.6
32.7
32.8
C6orf10
HLA-DRA
HLA-DRB1
BTNL2
HLA-DRB5
HLA-DQA1
C
Analysis of EBV-positive cHL
rs2734986
P = 1.2 x 10-15
rs6904029
29.6
29.7
29.8
29.9
30.0
30.1
30.2
30.3
HLA-F
HLA-G
HLA-A
ZNRD1
HCG9
PPP1R11
D
Analysis of EBV-negative cHL
rs6903608
P = 2.1 x 10-31
32.3
32.4
32.5
32.6
32.7
32.8
C6orf10
HLA-DRA
HLA-DRB1
BTNL2
HLA-DRB5
HLA-DQA1

Supplementary Figure 2. Plots of genome-wide associations between classical Hodgkin lymphoma (cHL) and single-nucleotide polymorphisms (SNPs) located within a 300 kb region flanking the index SNP (indicated by a black triangle) of the major histocompatibility complex region. The relationship between total cHL and two novel loci, **A**) rs2248462 (class I region, *MICB*) and **B**) rs2395185 (class II region, *HLA-DRA*), was investigated. **C)** The relationship between two loci in the *HLA-A* class I region indexed by rs2734986 and rs6904029 and Epstein Bar virus (EBV)-positive cHL was determined. **D)** The relationship between class II region locus, rs6903608 (*HLA-DRA*) and EBV-negative cHL was investigated. Multiple logistic regression was performed assuming a log-additive genetic model and adjusting for sex (male or female), country (France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. The $-\log_{10}$ (*P* value) for each SNP are plotted against their chromosomal position. All statistical tests were two-sided. The colors of the dots indicate the degree of linkage disequilibrium (based on r^2) in relation to the index SNP. Recombination rates (cM/Mb) overlay the plots and are based on HapMap Phase I and II data (http://hapmap.ncbi.nlm.nih.gov). cM/Mb = centiMorgans/megabase.



Supplementary Figure 3. Stratified and subgroup analyses of two single-nucleotide polymorphisms (SNPs) located outside the major histocompatibility complex region. The relationship between **A)** rs20541 (*IL13*) and **B)** rs27524 (*ERAP1*), with classical Hodgkin lymphoma (cHL) was investigated. Odds ratios (ORs), represented by boxes with the area of each box inversely proportional to the variance of the estimate) and 95% confidence intervals (CIs, **error bars**) were derived using multiple logistic regression assuming a

log-additive genetic model and adjusting for sex (male or female), country (up to eight indicator variables after excluding one country as the reference, depending on the analysis: France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors for the genome-wide association study (GWAS) analyses only. The **dashed vertical line** represents the OR of the SNP in the analysis of total cHL among all subjects and the width of the **diamond** is the corresponding 95% CI. The results are on the basis of a combined analysis of genome-wide association and independent replication phase results using inverse variance weighting meta-analysis. In the analysis stratified by study, the GWAS included the EPILYMPH study (EPILYMH-GWAS), Scandinavian Lymphoma Etiology Study (SCALE-GWAS), the Scotland and Newcastle Lymphoma Group and Young Adult Hodgkin Case-control Study analyzed together (referred to as the UK studies-GWAS), and the Northern Dutch Hodgkin Lymphoma Study (Netherlands-GWAS). Results by study for the independent replication included the EPILYMPH study (EPILYMPH-Replication), and the Scotland and Newcastle Lymphoma Group, Young Adult Hodgkin Case-control Study, and Epidemiology and Genetics Lymphoma Case-control Study analyzed together (referred to as the UK Studies-Replication). $P_{\text{homogeneity}}$ was on the basis of the Cochran Q test statistic and was used to evaluate between-study heterogeneity in results. Associations between the SNPs and cHL subgroups (including histologic subtype, EBV status, and age) were performed, and $P_{\text{homogeneity}}$ for the χ^2test of homogeneity indicates the differences in the OR between subgroup analyses. All statistical tests were two-sided. Ca = cases, Chr = chromosome, Co = control subjects, EBV = Epstein-Barr virus, MCHL = mixed cellularity Hodgkin lymphoma, NSHL = nodular sclerosis Hodgkin lymphoma.